

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 9, 2003, 16:11:58 ; Search time 8.22857 Seconds
(without alignments)
91.441 Million cell updates/sec

Title: US-09-905-691-4

Perfect score: 16

Sequence: 1 ARRAAARARRAEEA 16

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	7	43.8	263	1 CCMC_BRAJA	P30962 bradyrhizob
2	7	43.8	291	1 YM32_MYCTU	Q10515 mycobacteri
3	7	43.8	381	1 ARGJ_METH	P96137 thermus the
4	7	43.8	387	1 ARGJ_METH	O8tx15 m arginine
5	7	43.8	417	1 PROA_MEIRU	O86053 meliothermus
6	7	43.8	423	1 CES5_HUMAN	Q9brw7 homo sapien
7	7	43.8	428	1 Y486_MYCLE	P54138 mycobacteri
8	7	43.8	444	1 VGLX_HSVBS	Q08103 bovine herp
9	7	43.8	474	1 CYAE_BORPE	P11092 bordetella
10	7	43.8	480	1 Y486_MYCTU	Q11152 mycobacteri
11	7	43.8	521	1 EX7L_RHILO	Q987v3 rhizobium l
12	7	43.8	603	1 US26_HCMVA	P09699 human cytom
13	7	43.8	621	1 HEM1_AGABI	Q92403 agaricus bi
14	7	43.8	742	1 UL47_HSVBP	P30021 bovine herp
15	7	43.8	777	1 METE_CAUCR	Q9aaaw1 caulobacter
16	7	43.8	913	1 VGLB_PRVIF	P03179 epstein-bar
17	7	43.8	1318	1 VP14_EBV	P45387 haemophilus
18	7	43.8	1394	1 HAP_HAEIN	O15020 homo sapien
19	7	43.8	2390	1 SPCE_HUMAN	Q96dt5 homo sapien
20	7	43.8	4523	1 DIHE_HUMAN	P43740 haemophilus
21	6	37.5	68	1 RPOZ_NEIMA	Q9cmdb2 pasteurella
22	6	37.5	88	1 RPOZ_HAEIN	Q9kmn3 vibrio chol
23	6	37.5	89	1 RPOZ_PASMU	Q87cbo vibrio para
24	6	37.5	90	1 RPOZ_VIBCH	Q8ddv5 vibrio vuln
25	6	37.5	90	1 RPOZ_VIBPA	P08374 escherichia
26	6	37.5	90	1 RPOZ_VIBBU	Q8zjg3 versinia pe
27	6	37.5	91	1 RPOZ_ECOLI	Q10879 mycobacteri
28	6	37.5	91	1 RPOZ_YERPE	Q01182 rhodobacter
29	6	37.5	94	1 R28A_MYCTU	P16747 human cytom
30	6	37.5	108	1 NIEW_RHOSH	P44356 haemophilus
31	6	37.5	113	1 UL67_HCMVA	P36344 simian herp
32	6	37.5	117	1 RLJ8_HAEIN	
33	6	37.5	117	1 VGLJ_HSVBS	

RESULT 1

ID	CCMC_BRAJA	STANDARD;	PRT;	263 AA.
AC	P30962;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Heme exporter protein C (Cytochrome c-type biogenesis protein cycZ).			
GN	CYCZ OR CCMC OR BLR0469.			
OS	Bradyrhizobium japonicum.			
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;			
OC	Bradyrhizobiaceae; Bradyrhizobium.			
OX	NCBI_TaxID=375;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-110RIF15;			
RX	MEDLINE-91210304; PubMed-1850420;			
RA	Ramsel T.N., Winteler H.V., Hennecke H.;			
RT	"Discovery and sequence analysis of bacterial genes involved in the			
RT	biogenesis of c-type cytochromes.";			
RL	J. Biol. Chem. 266:7793-7803(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-USDA 110;			
RX	MEDLINE-22484998; PubMed-12597275;			
RA	Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,			
RA	Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,			
RA	Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,			
RA	Tabata S.;			
RT	"Complete genomic sequence of nitrogen-fixing symbiotic bacterium			
RT	Bradyrhizobium japonicum USDA110.";			
RL	DNA Res. 9:189-197(2002).			
CC	-1- FUNCTION: REQUIRED FOR THE EXPORT OF HEME TO THE PERIPLASM FOR THE			
CC	BIOTRANSFORMATION OF C-TYPE CYTOCHROMES.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane			
CC	(Probable).			
CC	-1- SIMILARITY: BELONGS TO THE CCMC/CYCZ/HELC FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; M60874; AAA36194.1;			
DR	EMBL; AF005936; BAC45734.1;			
DR	PIR; C39741; C39741.			
DR	InterPro; IPR002541; CytC_asm.			
DR	InterPro; IPR003557; CytC_blog_CcmC.			
DR	Pfam; PF01578; CytC_asm; 1.			
DR	PRINTS; PR01386; CCMCBIOGNIS.			
DR	TIGRFAM; TIGR01191; ccmC; 1.			
KW	Cytochrome c-type biogenesis; Transport; Transmembrane; Inner membrane; Complete proteome.			

Q9zh15 alcaligenes
Q8xv06 ralstonia s
P49926 canis famil
Q24709 synechococc
Q9hwe9 pseudomonas
P52767 pinus thunb
Q07178 rhodobacter
O53412 mycobacteri
Q33060 mycobacteri
Q10542 mycobacteri
Q9yfi92 aeropyrum p
Q9rz54 deinococcus

ALIGNMENTS

34 6 37.5 120 1 PAND_ALCEU
35 6 37.5 120 1 PAND_RALSO
36 6 37.5 122 1 CRF_CANFA
37 6 37.5 130 1 RS11_SYNPF
38 6 37.5 130 1 RS8_PSEAE
39 6 37.5 134 1 RK16_PINTH
40 6 37.5 135 1 NIUL_RHOCA
41 6 37.5 139 1 LPOV_MYCTU
42 6 37.5 143 1 Y880_MYCLE
43 6 37.5 143 1 Y880_MYCTU
44 6 37.5 143 1 RL32_AERPE
45 6 37.5 145 1 YV40_DEIRA

DR Pfam; PF00702; Hydrolase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 291 AA; 30694 MW; 750F090FB154E6E5 CRC64;

Query Match 43.89; Score 7; DB 1; Length 291;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RAAARRA 12
| | | | |
DB 66 RAAARRA 72

RESULT 3
ARGJ_THETH STANDARD; PRT; 381 AA.
AC P96137;
DT 15-SEP-2003 (Rel. 42; Created)
DT 15-SEP-2003 (Rel. 42; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Glutamate N-acetyltransferase (EC 2.3.1.35) (Ornithine
DE acetyltransferase) (Ornithine transacetylase) (OATase) (Contains:
DE Glutamate N-acetyltransferase alpha chain; Glutamate N-
DE acetyltransferase beta chain).
GN ARGJ.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB27;
RX MEDLINE=98154436; PubMed=9493385;
RA Baetens M., Legrain C., Boyen A., Glansdorff N.;
RT "Genes and enzymes of the acetyl cycle of arginine biosynthesis in the
RT extreme thermophilic bacterium thermophilus HB27.";
RL Microbiology 144:479-492(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HB27;
RC Sanchez R., Roovers M., Glansdorff N.;
RA "Organisation of arginine biosynthetic genes in Thermus
RT thermophilus.";
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RL -1- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate - L-
CC ornithine + N-acetyl-L-glutamate.
CC -1- PATHWAY: Arginine biosynthesis; fifth step.
CC -1- SUBUNIT: Heterotetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- MISCELLANEOUS: Some bacteria possess a monofunctional argJ, i.e.,
CC capable of catalyzing only the fifth step of the arginine
CC biosynthetic pathway
CC -1- SIMILARITY: Belongs to the argJ family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; Y10525; CAA71551.1; -;
DR EMBL; Y18353; CAA7143.1; -;
DR HAMAP; MF_01106; -; 1.
DR InterPro; IPR002813; ArgJ.
DR Pfam; PF01960; ArgJ; 1.
DR ProDom; PD004193; ArgJ; 1.
DR TIGRfam; TIGR00120; ArgJ; 1.
KW Arginine biosynthesis; transferase; Acyltransferase.
GN GLUTAMATE N-ACETYLTRANSFERASE ALPHA CHAIN
CHAIN 1 175
FT


```

Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 AAAAAAAR 10
Db 13 AAAAAAAR 19

RESULT 6
CES5 HUMAN STANDARD; PRT; 423 AA.
AC Q9BXW7; Q9BXW8; Q9NW48; Q9NX41;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
GN Cat eye syndrome critical region protein 5 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE-21275466; PubMed-11381032;
RA Footz T.K., Brinkman-Mills P., Banting G.S., Maier S.A., Riaz M.A.,
RA Brigland L.J., Hu S., Birren B., Minoshima S., Shimizu N., Pan H.,
RA Nguyen T., Fang F., Fu Y., Ray L., Wu H., Shaull S., Phan S., Yao Z.,
RA Chen F., Huan A., Hu P., Wang Q., Loh P., Qi S., Roe B.A.,
RA McDermid H.E.
RT *Analysis of the cat eye syndrome critical region in humans and the
RT region of conserved synteny in mice: a search for candidate genes at
RT or near the human chromosome 22 pericentromere.*;
RL Genome Res. 11:1053-1070(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Embryo, and Gastric carcinoma;
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Watanabe M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakematsu A., Nakamura Y., Nagahari K., Masubo Y.,
RA Ninomiya K., Iwayanagi T.;
RT *NEO human cDNA sequencing project.*;
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Lymph;
RX PubMed-12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schell J.E., Jones S.J.M., Marra M.A.;
RT *Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.*;
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- ALTERNATIVE PRODUCTS;
CC Event-Alternative splicing; Named isoforms-2;
CC Name-2;
CC IsoId-Q9BXW7-1; Sequence-Displayed;

```

```

Name-1;
IsoId-Q9BXW7-2; Sequence-VSP_003840;
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- MISCELLANEOUS: Candidate gene for the Cat Eye Syndrome (CES), a
CC developmental disorder associated with the duplication of a 2 Mb
CC region of 22q11.2. Duplication usually takes in the form of a
CC supernumerary bisatellited isodicentric chromosome, resulting in
CC four copies of the region (represents an inv dup(22)(q11)). CES is
CC characterized clinically by the combination of coloboma of the
CC iris and anal atresia with fistula, downslanting palpebral
CC fissures, preauricular tags and/or pits, frequent occurrence of
CC heart and renal malformations, and normal or near-normal mental
CC development.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF273271; AAK19152.1; -
CC EMBL; AF273270; AAK19151.1; -
CC EMBL; AK001034; BAA91475.1; -
CC EMBL; AK000461; BAA91180.1; -
CC EMBL; BC042540; ARA42540.1; -
CC Genew; HGNC:1843; CECR5.
CC InterPro; IPR006353; HAD_CECR5.
CC InterPro; IPR006357; HAD_SF_IIA.
CC TIGRFAMs; TIGR01456; CECR5; 1.
CC TIGRFAMs; TIGR01460; HAD-SF-IIA; 1.
CC Signal; Alternative splicing. POTENTIAL.
CC SIGNAL 1 23
CC CHAIN 24 423
CC CAT EYE SYNDROME CRITICAL REGION PROTEIN
CC VARSPLIC 1 41
CC -----
CC CONFLICT 54 54 D -> N (IN REF. 2; BAA91180).
CC CONFLICT 86 86 V -> F (IN REF. 2; BAA91475).
CC CONFLICT 297 297 E -> G (IN REF. 2; BAA91475).
CC SEQUENCE 423 AA; 46321 MW; C4D9208AB8B88CCE CRC64;
SQ
Query Match 43.8%; Score 7; DB 1; Length 423;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 RAARAAA 9
Db 19 RAARAAA 25
RESULT 7
Y486_MYCLE
ID Y486_MYCLE STANDARD; PRT; 428 AA.
AC P54138; Q9CB50;
DT 01-OCT-1995 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ML2443.
GN ML2443 OR U2168F OR B2168_C2_201.
OS Mycobacterium leprae
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;

```

```

RX MEDLINE-21128732; PubMed-11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Davlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Halroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- SIMILARITY: TO M.TUBERCULOSIS RV0486.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U00018; AAA17228.1; ALT_INIT.
CC EMBL; AL583925; CAC31960.1; -.
CC PIR; H87214; H87214.
CC Leproma; ML2443; -.
CC InterPro; IPR001296; Glyco_transf.1.
CC Pfam; PF00534; Glycos_transf.1; 1.
CC Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 428 AA; 45291 MW; A149F0U187E3587C CRC64;

Query Match 43.8%; Score 7; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRAARA 8
Db 204 RRAARA 210

RESULT 8
VGLX_HSVBS
ID VGLX_HSVBS STANDARD; PRT; 444 AA.
AC Q08103;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein GX precursor.
OC Bovine herpesvirus type 1.2 (strain ST).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=45407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94167875; PubMed-8122370;
RA Leung-Tack P., Audonnet J.F., Riviere M.;
RT "The complete DNA sequence and the genetic organization of the short
RT unique region (US) of the bovine herpesvirus type 1 (ST strain).";
RL Virology 195:409-421(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z23068; CAA80603.1; -.
CC PIR; S35783; S35783.
CC InterPro; IPR003363; Herpes_gg.
CC Pfam; PF02400; Herpes_gg; 1.
KW Glycoprotein; Transmembrane; Signal.

```

```

FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 444 GLYCOPROTEIN GX.
FT TRANSMEM 390 414 POTENTIAL.
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 444 AA; 46708 MW; 0145942AA35B05CB CRC64;

Query Match 43.8%; Score 7; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AARRARA 14
Db 422 AARRARA 428

RESULT 9
CYAE_BORPE
ID CYAE_BORPE STANDARD; PRT; 474 AA.
AC P11092;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein cyae precursor.
GN CYAE.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=18323;
RX MEDLINE-89091151; PubMed-2905265;
RA Glaser P., Sakamoto H., Bellalou J., Ullmann A., Danchin A.;
RT "Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-
RT haemolysin bifunctional protein of Bordetella pertussis.";
RL EMBO J. 7:3997-4004(1988).
CC -1- FUNCTION: CYAE IS NECESSARY FOR TRANSPORT OF CALMODULIN-SENSITIVE
CC ADENYLATE CYCLASE-HEMOLYSIN (CYCLOLYSIN).
CC -1- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PRTE FAMILY OF SECRETION PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X14199; CAA32414.1; -.
CC PIR; S02388; BVRCCE.
CC InterPro; IPR003423; OEP.
CC Pfam; PF02321; OEP; 2.
KW Hemolysis; Transport; Outer membrane; Signal.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 474 PROTEIN CYAE.
SQ SEQUENCE 474 AA; 50204 MW; 29A4F21D377FC957 CRC64;

Query Match 43.8%; Score 7; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AARRARA 14
Db 203 AARRARA 209

RESULT 10
Y486_MYCTU
ID Y486_MYCTU STANDARD; PRT; 480 AA.
AC Q11152;

```

```

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein RV0486
GN RV0486 OR MT0504 OR MTC2069.12.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=96344330;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigimeler K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtrovd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayan J.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uutterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO M. LEPRAE ML2443.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 277162; CAB00947.1; -.
DR EMBL; AE006951; AAK44727.1; -.
DR PIR; A70744; A70744.
DR TIGR; MT0504; -.
DR TubercuList; RV0486; -.
DR InterPro; IPR001296; Glyco_transf_1.
DR Pfam; PF00534; Glycos_transf_1; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 480 AA; 50541 MW; 2134755E994A9CCF CRC64;

Query Match 43.8%; Score 7; DB 1; Length 480;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRAARAA 8
DB 252 RRAARAA 258
|||||

RESULT 11
EX7L_RHILO STANDARD; PRT; 521 AA.
AC Q987V3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).

GN XSEA OR ML16896.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
CC -1- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -1- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003010; BAB3097.1; -.
DR HAMAP; MF_00378; -.
DR InterPro; IPR003753; Exonuc_VII_L.
DR InterPro; IPR004365; TRNA_anti.
DR Pfam; PF02601; Exonuc_VII_L; 1.
DR Pfam; PF01336; tRNA_anti; 1.
DR TIGR; TIGR00237; xsea; 1.
KW Hydrolase; Nuclease; Exonuclease; Complete proteome.
SQ SEQUENCE 521 AA; 56317 MW; 6FDE75D1A2DF085 CRC64;

Query Match 43.8%; Score 7; DB 1; Length 521;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ARAAAR 10
DB 306 ARAAAR 312
|||||

RESULT 12
US26_HCMVA STANDARD; PRT; 603 AA.
AC P09699;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical protein HHLF5.
GN US26.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87169717; PubMed=3031311;
RA Weston K., Barrell B.G.;
RT "Sequence of the short unique region, short repeats, and part of the
RT long repeats of human cytomegalovirus.";
```

```

RL J. Mol. Biol. 192:177-208(1986).
RN [2]
RP COMPLETE GENOME.
RX MEDLINE-90269039; PubMed-2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison S.A. III, Kouzarides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RT Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -1- SIMILARITY: BELONGS TO THE US22 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X17403; CAA35293.1; -
DR EMBL; X04650; CAB37117.1; -
DR PIR; A27216; QQBEDI.
DR InterPro; IPR003360; US22.
DR Pfam; PF02393; US22; 1.
KW Hypothetical protein.
FT CARBOHYD 44 44 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 603 AA; 70020 MW; E34F64D01E27687C CRC64;

Query Match 43.88; Score 7; DB 1; Length 603;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RRAARA 8
Db 574 RRAARA 580
|||||

RESULT 13
HEMI_AGABI STANDARD; PRT; 621 AA.
AC Q97403;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 5-aminolevulinic acid synthase, mitochondrial precursor (EC 2.3.1.37)
DE (Delta-aminolevulinic synthase) (Delta-ALA synthetase).
GN HEMI.
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_TaxID=5341;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D649;
RA Irgue E., Menak-Zunic M., Wood D.A., Thurston C.F.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + glycine -> 5-aminolevulinic acid + CO(2).
CC -1- COFACTOR: Pyridoxal phosphate.
CC -1- PATHWAY: Heme biosynthesis; first (rate-limiting) step.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; Z50096; CAA90424.1; -
DR InterPro; IPR003408; Ala_synthase.
DR InterPro; IPR004839; AminoTransf/2.
DR InterPro; IPR001917; NHtransf_2.
DR Pfam; PF02490; Ala_synthase; 1.
DR Pfam; PF00155; aminotran_1_2; 1.
DR PROSITE; PS00599; AA-TRANSFER-CLASS_2; 1.
KW Heme biosynthesis; Transferase; Acyltransferase; Mitochondrion;
KW Transit peptide; Pyridoxal phosphate.
FT TRANSIT 1 ? MITOCHONDRION.
FT CHAIN ? 621 5-AMINOLEVULINIC ACID SYNTHASE.
FT BINDING 362 362 PYRIDOXAL PHOSPHATE (PROBABLE).
SQ SEQUENCE 621 AA; 67426 MW; A335C3268FAE1AA3 CRC64;

Query Match 43.88; Score 7; DB 1; Length 621;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ARAAAR 10
Db 54 ARAAAR 60
|||||

RESULT 14
UL47_HSVBP STANDARD; PRT; 742 AA.
AC P30021;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE 80.7 kDa alpha trans-inducing protein (VP8 tegument protein) (VP7
DE protein) (107 kDa protein).
GN VP8
OS Bovine herpesvirus type 1 (strain P8-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10324;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92113550; PubMed-1662698;
RA Carpenter D.E., Misra V.;
RT "The most abundant protein in bovine herpes 1 virions is a homologue
RT of herpes simplex virus type 1 UL47."
RL J. Gen. Virol. 72:3077-3084(1991).
CC -1- FUNCTION: MODULATOR OF ALPHA-TIF (Vmw65 PHOSPHOPROTEIN) TRANS-
CC ACTIVATION. UL47 MAY HAVE KINASE ACTIVITY.
CC -1- SUBCELLULAR LOCATION: MAJOR TEGUMENT PROTEIN OF THE VIRIONS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE LATER STAGES OF INFECTION.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL47,
CC EBV-1 13, AND VZV 11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; D10327; BAA01170.1; -
DR EMBL; Z11610; CAA77683.1; -
DR PIR; JQ1435; TNBERI.
DR InterPro; IPR005029; Herpes_UL47.
DR Pfam; PF03362; Herpes_UL47; 1.
KW Transcription regulation; Trans-acting factor; Structural protein;
KW Late protein; Phosphorylation.
SQ SEQUENCE 742 AA; 80744 MW; 85979D8C2C953C89 CRC64;

Query Match 43.88; Score 7; DB 1; Length 742;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
OY 4 AAAAAA 10
DB 376 AAAAAA 382
|||||
DB 770 AAAAAA 776

Search completed: August 9, 2003, 16:29:51
Job time : 8.22857 secs

RESULT 15
METE_CAUOCR
ID METE_CAUOCR STANDARD; PRT; 777 AA.
AC Q9AAW1;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE 5-methyltetrahydropteroylglutamate--homocysteine methyltransferase
  (EC 2.1.1.14) (Methionine synthase, vitamin-B12 independent isozyme)
DE (Cobalamin-independent methionine synthase).
GN METE OR CC0482.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RA MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RA "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -|- FUNCTION: Catalyzes the transfer of a methyl group from 5-
CC methyltetrahydrofolate to homocysteine resulting in methionine
CC formation (By similarity).
CC -|- CATALYTIC ACTIVITY: 5-methyltetrahydropteroyltri-L-glutamate + L-
CC homocysteine = tetrahydropteroyltri-L-glutamate + L-methionine.
CC -|- COFACTOR: zinc; binds one ion per subunit (By similarity).
CC -|- PATHWAY: Terminal step in the de novo biosynthesis of methionine.
CC -|- SIMILARITY: Belongs to the vitamin-B12 independent methionine
CC synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE005721; AAK22469.1; -
CC PIR; A87309; A87309.
CC TIGR; CC0482; -.
CC HAMAP; MF_00172; -.
CC InterPro; IPR002629; Methionine_synth.
CC InterPro; IPR006276; Met_syn_B12ind.
CC Pfam; PF01717; Methionine_synth; 1.
CC ProDom; PD004692; Methionine_synth; 2.
CC TIGRFAMs; TIGR01371; met_syn_B12ind; 1.
CC Transferrase; Methyltransferase; Methionine biosynthesis; Zinc; Repeat;
CC Complete proteome.
CC METAL 665 665 ZINC (BY SIMILARITY).
CC METAL 667 667 ZINC (BY SIMILARITY).
CC METAL 750 750 ZINC (BY SIMILARITY).
CC SEQUENCE 777 AA; 84380 MW; 1C4D5FD7E80A80F3 CRC64;

Query Match 43.8%; Score 7; DB 1; Length 777;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 AAAAAA 14
```